



INSTITUTE FOR INFORMATION TECHNOLOGY APPLICATIONS

US AIR FORCE ACADEMY

Collaborative Tool for Command and Control Team Effectiveness Studies

Captain Chad Tossell, USAF

**Institute for Information Technology Applications
United States Air Force Academy, Colorado**

Technical Report TR-08 1

Approved for public release with unlimited distribution.

November 2008

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE NOV 2008		2. REPORT TYPE		3. DATES COVERED 00-00-2008 to 00-00-2008	
4. TITLE AND SUBTITLE Collaborative Tool for Command and Control Team Effectiveness Studies				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Information Technology Applications,United States Air Force Academy,United States Air Force Academy,CO				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 19	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Collaborative Tool for Command and Control Team Effectiveness Studies: Experimental Test of Interventions to Improve Performance in Command and Control

Table of Contents

Author's Biography	2
Abstract	3
Assessing the Effectiveness of Gaming Technology for Command & Control Training	4
Background	4
Method	4
Results	4
Collaboration Mediums, Team Performance, & Efficacy Perceptions	4
Background	4
Method	4
Results	5
Collaboration Methods in the Air & Space Operations Center (AOC)	5
Background	5
Method	7
Results	7
Research Assistants	7
Milestones	8
Final Results	8
Future Studies	8
References	8
List of Figures	
Figure 1 Cadets playing Airstrike game for study	5
Figure 2 Data Collection Window	7
List of Tables	
Table 1: Collaboration Tools and Team Processes	6
Attachments	
Attachment 1: Survey Form	9
Attachment 2: Conference Poster	10
Attachment 3: Conference Report	11
About the Institute	18

Biography

Captain Chad Tossell is an Instructor of Human Factors within the Behavioral Sciences and Leadership Department at the United States Air Force Academy. He directs the department's \$1.2 M research technology capabilities for faculty and cadets and the Cadet Summer Research Program. Captain Tossell graduated from the University of California, Berkeley, in 2003 and received his Masters of Science in Applied Psychology from Arizona State University.

The views expressed in this paper are those of the author and do not necessarily reflect the official policy or position of the Institute for Information Technology Application, the Department of the Air Force, the Department of Defense or the U.S. Government.

Comments pertaining to this report are invited and should be directed to:

Sharon Richardson
Director of Conferences and Publication
Institute for Information Technology Applications
HQ USAFA/DFPS
2354 Fairchild Drive, Suite 6L16D
USAF Academy CO 80840-6258
Tel. (719) 333-2746; Fax (719) 333-2945
<mailto:sharon.richardson@usafa.edu>

ABSTRACT

The Air and Space Operations Center (AOC) is the most complex weapons system in the United States Air Force (USAF) inventory (Garrity, Morley, Rodriguez, & Tossell, 2004). It is important for potential AOC operators to receive proper training in order to effectively accomplish the mission. This study had two goals. The first was to determine if an inexpensive, low fidelity strategy game, such as Airstrike, could be used to provide AOC familiarization training. The second was to experimentally test pathfinder network scaling methods in its ability to measure learning. Analysis of the gaming technology and SMARTboard showed that both were successful in improving team performance. Measuring knowledge structures using pathfinder was also shown to be a useful adjunct to performance-based methods of assessing training already employed.

Collaborative Tool for Command and Control Team Effectiveness Studies: Experimental Test of Interventions to Improve Performance in Command and Control

Assessing the Effectiveness of Gaming Technology for Command & Control Training

Background: The first study aimed at determining if an inexpensive, low fidelity strategy game, such as Airstrike, can be used to train potential AOC operators.

Method: We used pathfinder to assess how knowledge structures changed as a result of this game. Knowledge structures (i.e., mental models) represent how we organize concepts in our long-term memory. Research suggests that novices and experts arrange this information differently. Pathfinder software is a tool that can show representations of our conceptual mental models. We used a within-subjects design. Specifically, this study used 28 students with no prior AOC training and we presented AOC concepts to them in survey format. Cadets ranked the relevance of concepts to each other and pathfinder created knowledge structures from these scores. The students then completed game familiarization and training. After they completed this one-hour task, they completed the relevance rankings of the same survey of AOC concepts. These knowledge structures were compared to each other and showed significant differences after training had been completed. Knowledge structures were also obtained from subject matter experts (SMEs) via this same survey during a cadet summer research experience at CAOC-N, Nellis AFB, NV. SME reactions were also captured after they played the Airstrike game. The concepts and scale used in the survey are below shown in Attachment 1.

Results: As shown in the poster at Attachment 2, knowledge structures did change as a result of playing the Airstrike game. Additionally, post-training knowledge structures were significantly more coherent compared to pre-training knowledge structures ($p < .05$). Cadet knowledge structures also were more similar to referent structures obtained from SMEs. For a more thorough explanation of how knowledge structures are used to assess learning and how to interpret data such as above, please see the results section in the document that follows the report. Discussion: Thus, Airstrike Gaming Technology is a potentially effective way to provide familiarization training to future AOC Operators. SMEs thought this game might be helpful for trainees to think at the operational-level of warfare.

Collaboration Mediums, Team Performance, & Efficacy Perceptions

Background: In today's Air Force, the practice of distributed mission operations (DMO) requires teams to plan and execute complex missions even though they are unable to meet face to face. Many times these groups utilize synchronous and asynchronous communications to plan and execute missions.

Method: This study examines the impact of synchronous and asynchronous communications on team performance in the planning and execution of a mission. The study involved 2 x 2 mixed-factors ANOVA with a between-subjects factor of communication mode (2 levels) and a within-subjects factor of task (each group was measured twice). Groups in all conditions were given 15 minutes to plan for a team performance task with provided materials. They were then given 5 minutes to complete a team development survey. Then the groups were given 2 minutes to execute the team performance task and objective team performance metrics were measured. After a 10 minute break, the teams reconvened and were given another 15 minutes to plan a new design focusing on a similar, yet more difficult, team performance task. After planning, they were given 5 minutes to take another team development survey followed by a 2 minute execution phase. Team performance metrics were again measured after the second

execution phase. Fifty participants have completed the study to date, and approximately fifty more participants are currently completing the study.

Results: Preliminary results show: 1) that perceptions of teamwork and technical abilities are affected by the type and order of communication used, and 2) communication mode influences team performance. Please see Appendix 2 for slide presentation.



Figure 1 Cadets playing Airstrike game for study

Collaboration Methods in the Air & Space Operations Center (AOC)

Background: The Air and Space Operations Center (AOC) is the senior command and control (C2) node in the USAF Theater Air Control System. Consisting of a number of divisions, each comprised of multiple teams, the AOC provides operational level C2 of air and space forces as the focal point for planning, directing, and assessing air and space operations. Personnel assigned to the AOC are responsible for managing hundreds of aircraft and ground assets every day of an operation. Individuals working in an AOC must be well-trained, versed in military doctrine, and have a “big-picture” viewpoint of the given operation.

As complexity continues to grow in today’s war-fighting techniques, commanders and C2 operators in AOCs increasingly rely on technology to provide situational awareness. Specifically, researchers claim that “in the future, if soldiers are to function in a distributed fashion they will need collaborative tools and systems to exchange information and most importantly Situation Awareness (SA)” (Bolstad & Endsley, 2005). A common operating picture (COP) is one example of how a commander can maintain SA of the battlefield. The COP is used to track air, ground, and naval assets in an Area of Responsibility (AOR) while simultaneously performing intelligence, reconnaissance, and surveillance missions. The COP, in general, is a fairly new utility for the military and consequently has limited research available. The concept, however, is undergoing implementation across each branch of United States military service. For example, the Air Force uses a blue force tracker device to monitor the location of other airplanes, the Navy uses a COP to monitor the location of other ships, and the Army is implementing a COP in order to monitor friendly and foe forces on the ground within a given set of coordinates.

The purpose of the COP is to create a “single identical display of relevant information shared by more than one commander” (Pendall, 2005). In doing so, the COP acts as a method of communication that relays information to interested parties within the AOC. Furthermore, the COP provides a level of

situational awareness because the display shows the location of specifically defined objects (e.g. friendly aircraft, enemy aircraft, etc). Pendall also claimed that the COP offers an increased understanding about the target, which allows for a “faster decision cycle at all levels of command.” Although there are many positive aspects of the COP, there is the potential to rely too heavily on the display and focus more on planning reactively instead of proactively.

In addition to the COP, two important communication devices include face-to-face communication and chatting/instant messaging. As inferred from the name, face-to-face communication takes place between two or more people at the same location. There is no barrier to tone, expression, and other verbal communication characteristics. Chatting or instant messaging, on the other hand, is simply an exchange of words that occurs among technological devices. Chatting is severely limited in its ability as a communication device because no tone, expression, pictures, or lengthy messages can be shared.

The COP, face-to-face communication and chatting make up a set of collaboration tools which are used to “facilitate the communication and exchange of information among team members who are working together to complete a shared task” (Bolstad & Endsley, 2001). Among the three collaboration methods, each provide different levels of situational awareness (SA). The following chart developed by Bolstad and Endsley (2001) convey how specific collaboration tools aid team processes and provide different levels of shared situational awareness.

Table 1: Collaboration Tools and Team Processes

Tool Category	Planning	Brainstorming	Data Gathering	Data Distribution	Shared SA
Face-to-face	Good	Good	Moderate	Moderate	Medium-High
Video Conferencing	Moderate	Limited	Limited	Good	Medium-High
Audio Conferencing	Moderate	Limited	Limited	Good	Medium-High
Telephone	Moderate	Limited	Limited	Good	Medium-High
Net Radio	Moderate	Limited	Limited	Good	Medium-High
Chat/Instant Messaging	Poor	Poor	Limited	Moderate	Moderately Low
White Board	Moderate	Limited	Limited	Moderate	Moderate
File Transfer	Poor	Poor	Moderate	Moderately Good	Moderate
Program sharing	Moderate	?	Low	Low	Low
Email	Low	Poor	Low	Moderately Good	Moderately Low
Groupware	Poor	Moderate	Moderate	Low	Low
Bulletin Board	Poor	Moderate	Moderate	Low	Low
Doman Specific Tools	High	Limited	High	High	High

Victory Dialog

America Losses and Points			
Airplane Losses	12	Airplane Points	6580
Helicopter Losses	3	Helicopter Points	560
Ground Losses	0	Ground Points	0
SAM Site Losses	0	SAM Site Points	0
Recon Points	0	Rescue Points	0
No-Fly-Zone Points	0		

Cuba Losses and Points			
Airplane Losses	80	Airplane Points	15320
Helicopter Losses	0	Helicopter Points	0
Ground Losses	202	Ground Points	3600
SAM Site Losses	47	SAM Site Points	970
Recon Points	0	Rescue Points	0
No-Fly-Zone Points	0		

Victory Values			
Major Defeat	0	Minor Defeat	6000
Minor Victory	12000	Major Victory	18000

Total Points: 12750 Outcome: America Minor Victory

OK Help

Figure 2 Data Collection Window

As bolded above, face-to-face communication, chatting, and domain specific tools (COP) all have different levels of shared situational awareness with chatting having the lowest level of the three and the COP having the highest. However, while these various methods of communication have demonstrated a positive impact on SA, little research has been done to show how these methods directly affect team performance.

Method: The experiment was a between-subjects design. Participants in the study were divided into three groups and completed the same Command and Control task, each group using a different communication device. Our task was designed to require coordination, collaboration and action from both team members to achieve success. The first group performed a task using face-to-face communication (no COP or chat). The second group used an instant messaging program for team communication. Lastly, the third group used a Common Operating Picture (SMARTboard) for team coordination and communication. The team members in the second and third conditions were set-up in different rooms to prevent the use of face-to-face communication while using the COP or chatting devices. 36 cadets were used for this study. 12 cadets were assigned to each of the before mentioned groups, where each condition had 6 teams of 2 cadets each. Performance measures were captured from the game. We collected data from each of the items in this window.

Results: Teams that used a COP on the SMARTboard performed significantly better in many categories compared to the other two groups ($p < .05$). Text and face-to-face communication showed no differences in performance. In fact, cadets were very comfortable using chat to communicate during the scenario; this skill will likely transfer well to AOCs where chat is a primary means of communication. **Discussion:** COPs are indeed an effective way to display crucial information for team performance. Even in the absence of other communication methods, the COP was still the most successful method of collaboration. Clearly, the ability to include spatial information (i.e., a map with aircraft labels, directions and capabilities) in addition to verbal input helped groups perform well.

II. Research Assistants (number of semesters):

- | | |
|---------------------------|---|
| - Capt Ira Schurig (3) | - Ms. Teresa Bennett (4) – Capella University |
| - C1C Shaun Sucillon (2) | - C1C Jill Ward (1) |
| - C1C Anthony Rocco (1) | - C1C Benjamin Mendel (1) |
| - C1C Josh Splawn (2) | - C1C Adam Hood (1) |
| - C1C Brandon Wolf (1) | - C1C Raissa Kliatchko (1) |
| - C1C Jamie Moody (2) | - C1C Crystal Murray (1) |
| - C1C Daniel Kauffman (1) | - C1C Abby Barger (2) |
| - C1C Kamille Kemp (2) | |

All objectives were met at the appropriate times. Much of the groundwork was conducted before this project began with previous IITA support

[illegible]

IV. Final Results:

The results were provided to the Air Force Research Laboratory, Warfighter Readiness Research Division in Mesa, AZ. Additionally, cadets were able to participate in two conferences. Please see conference report at Attachment 3.

V. Future Studies

New researchers in the Department of Behavioral Sciences & Leadership plan to test the SMARTboard in the classroom. Additionally, results still unpublished from this study could be given at various conferences early next year.

VI. References

- Barger, A., Kemp, K., Tossell, C. (2007). Assessing Knowledge Acquisition in Academy Flight Screening. In Proceedings of the 29th Interservice/Industry Training Simulation & Education Conference, Orlando, FL.
- Tossell, C.C., Splawn, J.C., Denning, T.D., & Bennett, W.R. (2008). The Effectiveness of Gaming Technology in Command & Control Training. Presented at American Psychological Association Division 19 & Division 21 Annual Technical Symposium. Fairfax, VA.
- Schurig, I., Bennett, T., Sucillon, S.R., Moody, J.R., & Rocco, A.L. (2008). Collaboration Mediums, Team Performance, & Efficiency Perceptions Presented at American Psychological Association Division 19 & Division 21 Annual Technical Symposium. Fairfax, VA.

Survey

Please rate the relevance of each of the following concepts using the scale below. Please provide a rating for each cell (no blanks please).

0= The concepts have nothing to do with each other

1= The concepts have very little in common and are only slightly related

2=The concepts have one or two things in common but are still fairly separate

3=The concepts have some things in common and are moderately related to each other

4=The concepts have several things in common and are related

5=The concepts have a lot of things in common and are extremely related

How relevant are these...													
concepts...	...to these concepts?												
		A-10	ISR	EBO	F-16CJs	COG	Escort	SEAD	Weapon/ Aircraft Selection	Multiple Platform Integration	Target Prioritization	Dynamic Targeting	E-8C
A-10													
ISR													
EBO													
F-16CJs													
COG													
Escort													
SEAD													
Weapon/Aircraft Selection													
Multiple Platform Integration													
Target Prioritization													
Dynamic Targeting													
E-8C													

EBO: Effects Based Operations is focused on achieving desired results instead of simply selecting a target, because that type of target is generally useful.

ISR: Intelligence Surveillance and Reconnaissance deals with knowing the target locations, sizes, and capabilities. This is making sure all potential threats and targets are accounted for.

COG: Centers of Gravity are targets that will achieve desired effects if destroyed. An example of a COG would be an enemy Command Center. If leadership is taken out other functions are crippled or weakened.

Escort: Providing air support for bombers/CAS without air-to-air capabilities.

SEAD: Suppression of Enemy Air Defenses is making sure that enemy SAMs and AAA are defeated so that bombers and close air support can engage more easily targets.

Weapon/Aircraft Selection: Choosing the correct weapons system for a specific target or task.

Multiple Platform Integration: Coordinating attacks with different platforms.

Target Prioritization: Selecting appropriate targets first and adapting if new targets emerge.

Dynamic Targeting: Identifying and choosing platforms to defeat targets that emerge through increased intelligence or a change in the situation.



The Effectiveness of Gaming Technology in Command & Control (C2) Training



Cadet Joshua Splawn, Cadet Brandon Wolf, Capt Chad Tossell, USAF Academy

Purpose

Test the effectiveness of gaming technology to provide command & control familiarization training.

Air & Space Operations Center (AOC)

- Command and Control (C2) for air assets
- Mission: Plan, execute & assess air campaign in theater
- 5 Divisions
 - Strategy
 - Combat Plans
 - Combat Ops
 - ISR
 - AMD



Cadet Summer Research

- Combined Air Operations Center-Nellis (CAOC-N) at Nellis Air Force Base for 5 weeks
- Background research on gaming technologies
- Observed AOC operations
- Became familiar with Airstrike Game
- Collected reaction data from subject matter experts (SMEs)

Game—Airstrike

- Capabilities
 - Air Tasking Order Development
 - Combat Ops
 - Team-based
 - Customizable scenarios and flight plans



- Limitation
 - Differences between AOC tools & game

Methods

- Kirkpatrick's Model of Evaluating Training Effectiveness
- Within subjects design
 - Participants filled out survey relating AOC concepts
 - Receive training on Airstrike
 - Performed simple, custom scenario
 - Filled out same survey after performing scenario



Concepts

- Examples of Concept Descriptions
 - Effects Based Operations:**
Is focused on achieving desired results instead of simply selecting a target, because that type of target is useful
 - Centers of Gravity:**
Targets of great importance that will achieve desired effects if destroyed.

A-10
Intelligence, Surveillance, Reconnaissance (ISR)
Effects Based Operations (EBO)
F-16 Ch
Centers of Gravity (COG)
Escort
Suppression of Enemy Air Defenses
Weapon/Aircraft Selection
Multiple Platform Integration
Target Prioritization
Dynamic Targeting
E-3C

Measuring Knowledge Acquisition

- Pathfinder
 - Uses pairwise ratings between concepts to develop knowledge structures
 - Provides quantitative differences before and after training intervention
 - SME assessment of knowledge structure differences

Initial Results

- SME Assessment
Game helped participants better relate certain concepts [i.e., Centers of Gravity (COG) with Effects Based Operations (EBO) & Intelligence (ISR) with Weapon/Aircraft Selection]



SME Reactions

- Pros
 - Inexpensive, low fidelity strategy game
 - Developed to accurately reflect flight characteristics/capabilities of a wide variety of aircraft
 - May provide good theoretical foundation for familiarization training
- Cons
 - No specific skills transferable to AOC planners
 - Simplistic
 - Few similarities with team processes in AOC

Future Research & Development

- SME feedback to detail additional functions for Airstrike
 - MAAP toolkit
 - Planning sheets
 - Chat tool
- Additional SME Reactions
 - Additional data collection

Assessing Knowledge Acquisition in Academy Flight Screening (AFS)

Lt Abby Barger, Lt Kamille Kemp
Department of Behavioral Sciences & Leadership
USAF Academy, CO
kamille.kemp@wpafb.af.mil

Capt Chad Tossell
Department of Behavioral Sciences & Leadership
USAF Academy, CO
chad.tossell@usafa.edu

INTRODUCTION

Academy Flight Screening

The AFS program at the United States Air Force Academy aims to screen, motivate, and train potential pilot candidates. Most cadets participate in this program during the academic year. Consequently, students are also taking 12-18 hours of additional college credit, participating in military training, and involved with some form of organized athletic activity (some even at the intercollegiate level). The entire AFS course is 47.5 hours. It is almost equally divided between academic/ground school and flying training. Before a student is allowed to attend undergraduate pilot training (UPT), he or she must successfully complete the AFS program by demonstrating the ability to satisfactorily fly an aircraft and passing a written examination.

Assessment of Flight Training

The central goal of pilot training at any level within the USAF is to develop expertise in flight performance. Typical methods to assess the effectiveness of this training may include evaluations of changes in knowledge as a function of training and performance results as observed and rated by instructor pilots. In this project, we concentrated solely on assessing knowledge acquisition outside of the live training environment. Specifically, we chose to approach measuring the expertise of AFS graduates by examining how they conceptually organize various aspects of basic flight and comparing this knowledge structure with non-AFS graduates.

Researchers have utilized several methods for researching and defining knowledge structures. For example, multidimensional scaling (MDS) encompasses a variety of techniques that attempt to find a spatial layout of conceptual memory. MDS can represent large amounts of data in an easy to interpret format but it fails to capture asymmetrical and local relationships between concepts (McDonald and Schvanaveldt, 1988). Alternatively, discrete models use graphs with nodes (concepts) and links (distances between concepts) to

describe the local networking of concepts in memory. Pathfinder is an example of this sort of technique.

Pathfinder is a tool that transforms judgments of the relationship between concepts made by subjects into a corresponding knowledge structure. Other labels that may also be attached to this output include mental model, concept map, cognitive structure, and network structure. Generally, knowledge structures reflect the user's understanding of a system and can be created spontaneously or taught through training. Creating a correct mental model through explicit training is advantageous because the students are then able to learn more effectively and adapt better to novel situations (Wickens & Hollands, 2000). The study of mental models has many applications especially in the realm of education and training. For instance, if novices learn not only the information that they need to know but also an effective way of organizing it, they will more quickly become experts. Our objective in this research is to determine if AFS students are developing a specific and effective mental model of basic flight concepts.

Multiple research efforts highlight the difference in knowledge structures of experts versus novices in various domains. Schvanaveldt, Durso, Goldsmith, Breen, and Cooke (1985) addressed the differences between memory structures of expert and novice Air Force pilots. Cooke and Schvanaveldt (1988) conducted two studies exploring differences in cognitive structures between novice and experts in the field of computer programming. One conclusion reached from these research efforts is that not only do experts possess more knowledge than novices, but knowledge that is shared by the two groups is organized differently in memory (Cooke & Schvanaveldt, 1988).

The purpose of this research is to determine if there are measurable differences between the knowledge structures of AFS students and cadets who have not been through the AFS program. There are two hypotheses for this study. First, there will be a difference in knowledge structures of basic flight concepts between those cadets who have completed AFS (more experienced) and those who have not

(novices), indicating the training effectiveness of the AFS program. Second, the more experienced group will rate concepts more similarly than the novice group. This study also has the potential to allow the AFS program to track progress of students in a more objective way which will allow them to more effectively train future Air Force Pilots.

METHOD

Participants

A total of 36 USAFA cadets participated in this study, 18 of those had completed AFS and 18 had not. All participants were students enrolled in an Introductory Leadership Course and responded to a recruitment flyer. Demographic data was collected via a short survey at the beginning of the study. Subjects varied very little in flying experience and their objectives for completing AFS.

Materials

Two Dell Latitude D600 laptop computers were used to administer a survey via a graphical user interface created with MATLAB programming software. Pathfinder software, also MATLAB-based, was used for analysis of the data.

Procedure

Participants were recruited via the participant pool managed by the Behavioral Sciences & Leadership Department at the USAF Academy. On the recruitment flyer, every cadet indicated whether or not s/he had completed the AFS program. This information was used to sort the participants into either the non-AFS (novice) group or the AFS (more experienced) group. Appointments were scheduled individually through email. Upon arrival at the Behavioral Sciences Laboratory, participants were pre-briefed, informed of

Rate the relevance between the following two concepts:

ground_track

☐ 0 - Not Relevant

☒ 1 - Of Little Relevance

☐ 2 - Of Some Relevance

☐ 3 - Moderately Relevant

☐ 4 - Very Relevant

☐ 5 - Extremely Relevant

engine_stoppage

Next

EXPLANATION OF THE RATING SCALE

0= The concepts have nothing to do with each other

1= The concepts have very little in common and are only slightly related

2=The concepts have one or two things in common but are still fairly separate

3=The concepts have some things in common and are moderately related to each other

4=The concepts have several things in common and are related

5=The concepts have a lot of things in common and are extremely related

Figure 1: Questionnaire Screen

the risks inherent in the study and asked to sign an informed consent document.

The participants were first asked to provide demographic data on items that could potentially affect their knowledge of the flight concepts presented during the study and/or their motivation to learn flight concepts. It asked whether they had taken AFS or not, whether they were planning on taking AFS, whether they had a private pilot's license and whether or not they had taken Aeronautical Engineering 315 (a core class that teaches basic aerodynamic principles). Upon completing the demographic data, participants interacted with the screen as shown in Figure 1. This was the questionnaire that asked users to rate the relevance between a pair of concepts. When all pairs had been rated, the survey program output the participant data into text files which could be read by Pathfinder software.

After completion of the demographic survey, each participant completed the basic flight concepts survey. There were 30 flight concepts (Table 1) presented pairwise for a total of 435 questions.

No pairs were repeated or swapped, for example, a

Table 1. Flight Concepts

traffic pattern	runway orientation
coordinated flight	rudder
turn and slip indicator	yaw
proper configuration	power on stall
flaps	pitch
trim	power off stall
boldface	Vr
emergency procedures	steep approach
forward slip	SFL
ground track	engine stoppage
airspeed	glide speed
aim point	checklist
landing	proper spacing
take off	breakout procedures
wind conditions	centerline

participant would not be asked to rate both 'trim vs. flaps' and 'flaps vs. trim'. For each pair, the participant was required to rank the relevance between the two concepts on a scale from 0-5, with 0 indicating no relevance between the concepts and 5 indicating maximum relevance between the concepts. A key for the scale appeared on the screen for reference.

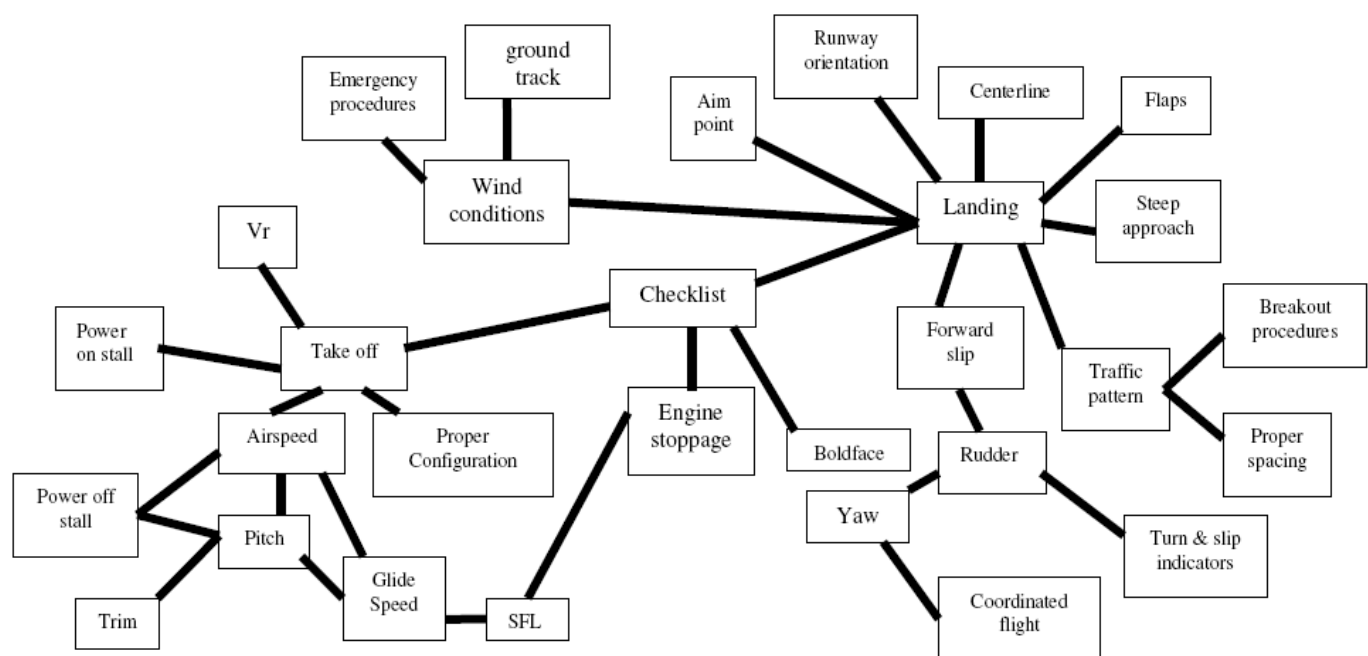


Figure 2. The Pathfinder Network Derived from the Average of AFS Graduate Ratings.

After completion of the survey, the facilitator ensured that the data file had properly saved each participant's responses to the survey. Participants were debriefed and given an extra credit receipt. The entire time spent in the laboratory ranged from 45 to 90 minutes.

All variables used in data analysis were based on the cadets' ratings of all the pairs of concepts. The resulting matrix after completion of survey contained 435 pairwise similarity measures representing proximities among all pairs of the 30 basic flight

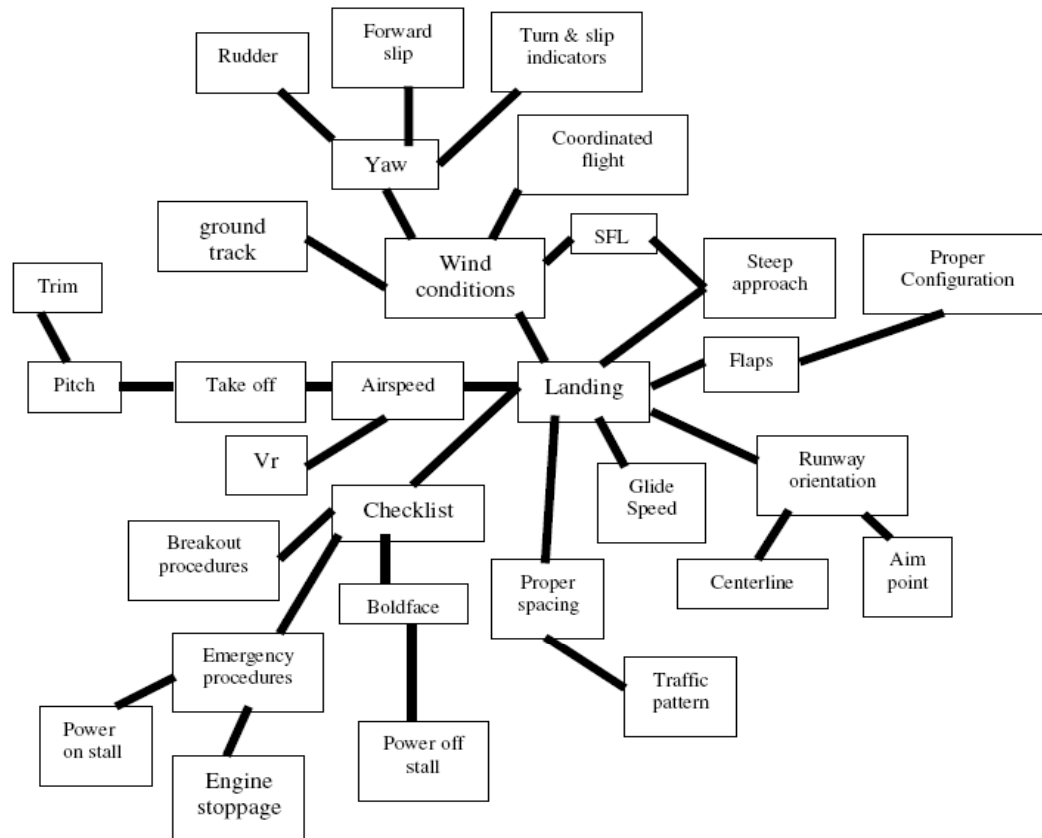


Figure 3. The Pathfinder Network Derived from the Average of Novice Ratings.

concepts. Such a matrix was obtained for each subject. The calculated psychological distances between the concepts were measures of their similarities to each other in terms of their relevance. The larger numbers are of greater similarity than the smaller numbers. The resulting matrix is a table in which both the rows and columns are the units of analysis and the cell entries are a measure of distance for each pair of cases. The difference between sets of qualitative elements (features) here forms the basis of a similarity model that can be used for fitting a network.

A network was derived from each set of ratings. Several additional variables were calculated. First, the data yielded coherence measures which assess the internal consistency of the ratings. Coherence is computed in two steps. First, a correlation of ratings for each pair of items is computed. For example, for

items 1 and 2, the ratings of item 1 across all of the other items is correlated with item 2 across all of the other items. This determines the extent to which the items in a pair are rated similarly against the other items. The second step correlates these correlations with the ratings given for the pairs. This can be seen as comparing the direct rating of a pair with the indirect relatedness inferred from the similarity of the ratings for the items in each pair. The coherence measure has been shown to reflect levels of expertise in that raters with more expertise in the concept domain generally produce higher coherence scores compared with less experienced raters. Coherence can also reveal unsystematic ratings that might be provided by a participant who does not take the rating task seriously resulting in careless ratings. Given that the rating task requires a large number of judgments, it is difficult for participants to be consistent by remembering earlier ratings. Rather, consistency

more likely stems from using a clear understanding of the domain as a basis of the ratings.

We also defined variables to assess the extent to which the networks obtained from individuals were similar to each other. Network similarity is assessed by first computing the ratio of the number of links two networks have in common over the number of unique links in both networks (i.e. the ratio of the cardinality of the intersection of the links in the two networks over the cardinality of the union). In a second step, the expected value of this ratio is subtracted from the ratio to yield the difference between the obtained value and the value expected by chance.

RESULTS

Pathfinder organized the data in a network consisting of 30 nodes representing the 30 basic flight concepts. The nodes were connected with weighted links based on the similarity data. Both the experienced (AFS graduates) and novice (non-AFS graduates) networks appear in figures 2 and 3. These connections represent the aggregated relatedness ratings between concepts as rated by both experienced and novice groups.

The novice structure was only a bit more complex (46 links) compared to the expert structure (33 links). Pathfinder also showed differences between experts and novices in terms of network similarity. This measure accounts for the number of shared links and intersections between individual networks taking into account network size and corrected for error. Two of the same networks would have a network similarity rating of 1. Two completely different networks (no similarities) would have a rating of 0.

In this study, both AFS graduates and novices had relatively low network similarity results. AFS graduates shared only 30% of their organization of basic flight concepts. Novices, on the other hand, shared an even fewer 20%. Thus, networks produced from the AFS graduate data ($M = .05$, $SD = .0117$) were significantly more similar to each other compared to the novices' networks ($M = .004$, $SD = .0113$), $t(14) = 5.86$, $p < .05$. Experienced respondents more consistently rated items in various combinations in the matrix and shared more characteristics in their conceptual structures compared to novices. Pathfinder's ability to classify experts and novices here is considered one

method to validate the precision of the network (Schvaneveldt et al, 1985).

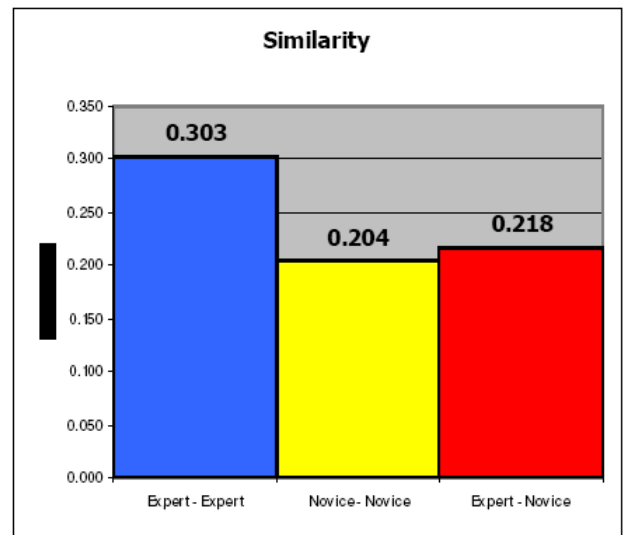


Figure 4. Network Similarity

Number of links in common is a measure of how similar two groups of networks are to each other. Figure 4 shows that the AFS networks had more links in common with each other than they had in common with the novices and than the novices has in common with the other novices.

The novices had about the same number of links in common with other novices as they did with those more experienced. The average novice network had only 35 links in common with the average expert network. The more experienced subjects on the other hand typically had approximately 45 links in common with each other. The probability that the more experienced networks would have this many links in common due merely to chance is .08%. Correcting for chance we find that the more experienced subjects still share approximately 23 links while the novices share only 12 and the two groups have about 13 in common. There is a significant difference between the number of links in common amongst the more experienced group as compared to the novice group $t(29) = 5.87$.

Identifying Discriminating Concepts

A final analysis was focused on identifying the concepts that are viewed most distinctly by experienced and novice pilots. Ratings of novice pilots were averaged. Similarly, ratings of AFS graduates were averaged as well. The differences

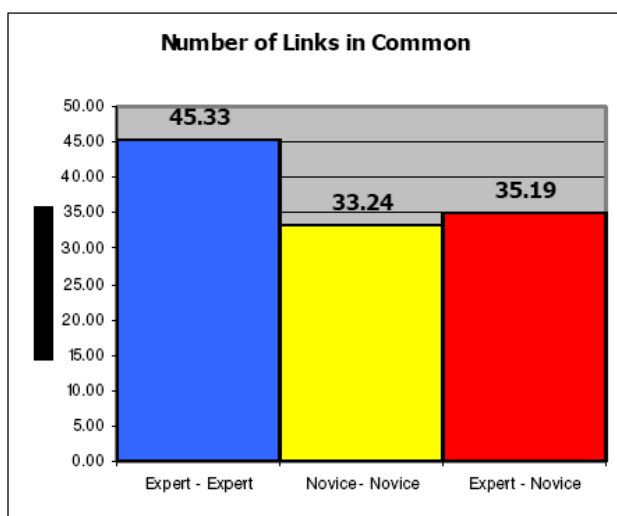


Figure 5: Number of Links in Common

between these two scores were computed. Summing the absolute values of these differences for each of the 30 concepts allows us to order them by the magnitude of the differences between experienced and novice ratings. This ordering is shown in Table 2.

Table 2. Concepts that Discriminate

Most Discriminating		Least Discriminating	
Rank	Concept	Rank	Concept
1	Emergency Procedures	1	Trim
2	SFL	2	Pitch
3	Proper configuration	3	Yaw
4	Power off stall	4	Turn & Slip Ind.
5	Coordinated flight	5	Rudder
		6	Boldface
		7	Runway Orientation
		8	Landing
		9	Aim Point

Knowing about the relative difficulty of these concepts could usefully feed back into the design of training by showing what the novices seem to know the least about. By ensuring that these concepts receive adequate coverage in the briefing and debriefing sessions, novices may be better prepared to deal with the corresponding aspects of their AFS flights.

In a qualitative comparison of the network diagrams produced by members of each group, there appears to

be only a slight difference. The simulated forced landing node in particular is telling. Our AFS group linked engine stoppage (the general cause of a forced landing) and glide speed (the most important consideration when making a forced landing) as most similar. Our novice group rated steep approach (a type of landing approach that may or may not be used in a forced landing) and wind conditions (an important consideration whenever landing an airplane) as most similar. We can see that although our novice group shows some understanding of basic aerodynamics they do not grasp some of the finer points of flying.

One potential reason for the lack of qualitative differences in networks could be that most cadets that participated in this study took an academic course on the fundamentals of aeronautics. It is possible that a set of more advanced concepts would reveal bigger differences. In addition, other measurement methods that focus on differences in procedural knowledge could be pursued as well.

DISCUSSION & CONCLUSIONS

Several aspects of the results reported above show that AFS training does produce a measurable difference in the way cadets conceptually organize basic flight information. In addition, there were overall increases in similarity among pilots after training. The increased similarity is apparently due to pilots being exposed to similar material/scenarios in training rather than to pilots learning more about members of their teams. Other studies of team cognition (see Cooke, Salas, Cannon-Bowers, & Stout, 2000) have often looked at increasing similarity among team members. The data reported here suggest that it may be useful to determine the extent to which increasing similarity is due to exposure to similar task conditions as opposed to training experiences with members of a team.

The results of our study could have some effects on the design of AFS training. The concepts we identified as leading to the greatest differences between novices and AFS graduates could receive more attention in the briefing and debriefing sessions accompanying the training sessions. Other ways in which the information we collect can be used in training are being considered. For example, students could be presented with networks like the one shown in Figure 2 to elicit discussion about the nature of the relationships depicted by the links. Inevitably such discussion draw out similarities and differences in the ways students think about the concepts. A hypothesis

that could be investigated is that training would benefit from such discussions, discussions that are driven by pilots underlying conception of flight.

Lastly, above results also show the sensitivity of pathfinder in capturing small differences in the representation of knowledge structures between slightly more experienced and novice individuals. AFS is primarily intended to screen potential flight candidates and only secondarily designed to provide basic flight training. Pathfinder's ability to provide quantitative distinctions in performance within the more experienced-novice paradigm here support future research into the utility of this tool as an adjunct to other performance measurement tools being developed by the Air Force Research Laboratories.

REFERENCES

- Brannick, M., Prince, C., & Salas, E. (2005). Can PC-Based Systems Enhance Teamwork in the Cockpit? In International Journal of Aviation Psychology (pp.173-187). Lawrence Erlbaum Associates, Inc.
- Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. (2000). *Measuring team knowledge*. Human Factors, 42, 151-173.
- Cooke, N. & Schvaneveldt, R. (1988). Effects of computer programming experience on network representations of abstract programming concepts. Int. J. Man-Machine Studies, 407-427.
- Gillan, D. J., & Schvaneveldt, R. W. (1999). Applying cognitive psychology: Bridging the gulf between basic research and cognitive artifacts. In F. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, M. Chi, & S. Lindsay (Eds.), *The handbook of applied cognition*. (pp. 3-31) Chichester: Wiley.
- Goldsmith, T. E., Johnson, P. J., Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83, 88-96.
- Hinkle, D., Wiersma, W., & Jurs, S. (1994). Applied statistics for the behavioral sciences (3rd Ed.). Princeton, NJ: Houghton-Mifflin.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Upper Saddle River, NJ: Prentice Hall.
- McDonald, J. & Schvaneveldt, R. (1988). The application of user knowledge to interface design. 289-338. In R. Guindon (Ed.), Cognitive Science and its Applications for Human-Computer Interaction (pp. 289-338). Hillsdale, NJ: Erlbaum.
- Rowe, A. L., Cooke, N. J., Hall, E. P. & Halgren, T. L. (1996). Toward an online assessment methodology: Building on the relationship between knowing and doing. *Journal of Experimental Psychology: Applied*, 2, 31-47.
- Schvaneveldt, R.W., Durso, D.T., Goldsmith, T.E., Breen, T.J, Cooke, N.M. (1985). Measuring the structure of expertise. Int. J. Man-Machine Studies, 699-728.
- Schvaneveldt, R., Durso, R., & Dearholt, D. (1989). Network structures in proximity data. The Psychology of Learning and Motivation, 249-284. Academic Press, Inc.
- Tossell, C. (2006). Representing the Structure of Expertise using Competency Constructs. Unpublished Thesis, Arizona State University.
- Wickens, C.D., & Hollands, J.G. (2000). Engineering Psychology and Human Performance. Upper Saddle River, NJ: Prentice-Hall Inc.

About the Institute

The Institute for Information Technology Applications (IITA) was formed in 1998 to provide a means to research and investigate new applications of information technology. The Institute encourages research in education and applications of the technology to Air Force problems that have a policy, management, or military importance. Research grants enhance professional development of researchers by providing opportunities to work on actual problems and to develop a professional network.

Sponsorship for the Institute is provided by the Assistant Secretary of the Air Force (Acquisition), the Air Force Office of Scientific Research, and the Dean of Faculty at the U.S. Air Force Academy. IITA coordinates a multidisciplinary approach to research that incorporates a wide variety of skills with cost-effective methods to achieve significant results. Proposals from the military and academic communities may be submitted at any time since awards are made on a rolling basis. Researchers have access to a highly flexible laboratory with broad bandwidth and diverse computing platforms.

To explore multifaceted topics, the Institute hosts single-theme conferences to encourage debate and discussion on issues facing the academic and military components of the nation. More narrowly focused workshops encourage policy discussion and potential solutions. IITA distributes conference proceedings and other publications nation-wide to those interested or affected by the subject matter.